



PCT

## RAW SEQUENCE LISTING

DATE: 10/08/2003

PATENT APPLICATION: US/10/088,117A

TIME: 14:51:42

Input Set : A:\05081004.app

Output Set: N:\CRF4\10082003\J088117A.raw

3 <110> APPLICANT: GIORDANO, GERARD  
 4 DOS SANTOS, JEAN-PHILIPPE  
 5 MEJEAN, VINCENT  
 7 <120> TITLE OF INVENTION: NUCLEOTIDE SEQUENCES DERIVED FROM GENES CODING FOR  
 8 TRIMETHYLAMINE N-OXIDE REDUCTASE, AND USES THEREOF,  
 9 ESPECIALLY FOR THE DETECTION OF BACTERIA  
 11 <130> FILE REFERENCE: 0508-1004  
 13 <140> CURRENT APPLICATION NUMBER: 10/088,117A  
 C--> 14 <141> CURRENT FILING DATE: 2002-09-30  
 16 <150> PRIOR APPLICATION NUMBER: PCT/FR00/02578  
 17 <151> PRIOR FILING DATE: 2000-09-15  
 19 <150> PRIOR APPLICATION NUMBER: FR 99/11543  
 20 <151> PRIOR FILING DATE: 1999-09-15  
 22 <160> NUMBER OF SEQ ID NOS: 43  
 24 <170> SOFTWARE: PatentIn Ver. 2.1  
 26 <210> SEQ ID NO: 1  
 27 <211> LENGTH: 2487  
 28 <212> TYPE: DNA  
 29 <213> ORGANISM: Shewanella massilia  
 31 <400> SEQUENCE: 1  
 32 atgaacagaa gagacttttt aaagggtatc gcctcactct ctttcgttgt cttaggtggc 60  
 33 agctcagtgt taacgccctt aaatgcctta gccaaagcgg gcatcaatga agatgaatgg 120  
 34 ctaaccacag gttcacactt cggcgccttt aaaatgaagc gcaaaaacgg cgtcattggc 180  
 35 gaagtgaaac cttcgcactt agataagtat ccaacggata tgattaacgg catccgcggc 240  
 36 atggtctaca atccatcgcg tgtacgttac cctatggtgc gcttagattt ttactcaaa 300  
 37 ggtcataaga gtaataccca tcaacggggt gatttccgct ttgttcgctg aacgtgggac 360  
 38 aaggcattaa cactgtttta gcattcatta gatgaagtcc aaacccaata cgggccatca 420  
 39 ggtctgcatg cggggcaaac cgggttgccg gccactggtc aactgcattc cagcacgagt 480  
 40 catatgcaac gtgcggtggg gatgcacggc aactatgtta agaaaatcgg cgactactcc 540  
 41 acaggtgcag gccaaactat tctgccctac gtgttaggtt caaccgaagt gtatgccag 600  
 42 ggcacttcat ggccgctgat cttagaacac agcgacacta tcgtgctgtg gtcgaacgat 660  
 43 ccgtacaaga acctgcaagt ggggttgaat gcggaacccc atgaatcttt tgcttatctt 720  
 44 gcgcagttaa aagagaaagt gaagcaaggc aagatccgcg tgatcagtat cgaccctgtg 780  
 45 gtgactaaga cccaagccta tttgggctgc gagcaactct acgttaaccc acagacagat 840  
 46 gtgaccttaa tgctggccat cgcccacgag atgatcagca cctatgtgat gggactataa 960  
 47 tttatccaag gctacagctt aggttttgaa gagtttggc cctatgtgat gggactataa 1020  
 48 gatggcgtag ccaaaacccc agaatgggccc gcgcctatct gtggtgttga agcccatgtt 1080  
 49 atccgcgact tggctaaaac cttagtcaag ggccgcactc agttcatgat gggctggtgt 1140  
 50 atccagcgcc agcaacacgg cgaacaaccc tattggatgc cggcggtact ggcgacctg 1200  
 51 atcggccaaa tcggtctacc cgggtggggt atcagctatg gtcaccacta ctcgagtatc 1260  
 52 ggcgtgcctt catcgggtgc cgctgcgcca ggtgccttcc cccgtaactt ggacgaaaat 1320  
 53 caaaagccac tctttgatag ctcagacttc aaggcgcgca gtagcacgat tccggttgcc 1380  
 54 cgtggtgattg atgcgattct cgaaccgggt aaaaccattg atgctaacgg ctcgaaagtg

ENTERED

7.6

**RAW SEQUENCE LISTING**  
**PATENT APPLICATION: US/10/088,117A**

DATE: 10/08/2003  
 TIME: 14:51:42

Input Set : A:\05081004.app  
 Output Set: N:\CRF4\10082003\J088117A.raw

```

55 gtttatcccg atatcaagat gatgattttc tcgggtaata atccttgga ccatcaccaa 1440
56 gacagaaacc gcatgaagca agccttccat aagcttgagt gtgtgggtcac tgtcgatgtg 1500
57 aactggacgg caacttgccg cttctcggat atcgtactgc ccgcttgtag tacctatgag 1560
58 cgcaacgata tcgacgtgta cggcgccctat gctaaccgcg gtatttttagc catgcagaaa 1620
59 atgggttgagc cactgtttga tagcttgctg gattttgaaa ttttctactcg ctttgccgcc 1680
60 gtactcggca aagagaaaga atacaccctg aacatgggagc aaatggagtgt gttagaaacc 1740
61 ctctataacg aatgtaaagc cgccaacgcg ggcaagtttg agatgcctga ctttgcgact 1800
62 ttctggaac aaggttatgt gcatttttgg gacggtgaag tctggacgcg ccatgcagac 1860
63 ttttagaaacg atcctgaaat caatccacta ggacgcctt caggtttgat tgaaatcttt 1920
64 agccgtaaga ttgatcaatt cggttacgat gactgtaaag gtcaccaaac gtggatggag 1980
65 aaaaccgagc gtagtcatgg cggccctggc tctgacaagc atccgatttg gttgcagtca 2040
66 tgccatccag acaaacgttt acactcgag atgtgtgagt cgcgagaata ccgcgagact 2100
67 tacgcagtca atggccgtga gctgtgtat atcagccctg tcgacgcaaa agcacgtggc 2160
68 atcaaagatg gcgatatagt gcgagcttt aacgaccgtg gccaaactgtt ggccgggtgct 2220
69 gtggtatcgg acaacttccc taaagggatt gtgcgaattc acgaaggcgc gtggtatggg 2280
70 ccagtaggta aagatggtag cactgaaggt ggtgctgaag tcggcgccct gtgtagttat 2340
71 ggcgatccta acaccctcac ttttagacata ggcacctcta aacttgccca agcttgctca 2400
72 gcctatactt gcttagtcga gtttgaaaaa taccaaggca aagtgcctaa ggtcagctcc 2460
73 ttcgatggcc cgatcgaagt cgaaatc
76 <210> SEQ ID NO: 2
77 <211> LENGTH: 2486
78 <212> TYPE: DNA
79 <213> ORGANISM: Shewanella putrefaciens
81 <400> SEQUENCE: 2
82 atgaacagaa gagacttttt aaaaggctta gcctcaacct ctttcgttgc ttttaggtggc 60
83 agctcagtac tagcgccctt aaatgcgctg gccataactg gcctgaatga aaacgaatgg 120
84 ctgaccactg gctcccactt cgggtgccttt aaaatcaagc gtaaaaacgg catgattgcc 180
85 gaagtcaaaag ctttcgattt agataaatat ccaacggata tgattaacgg tatccggggg 240
86 atggtctata acccatcccg cgtgcgttac ccgatggttc gcttagactt tttactaaaa 300
87 ggccataaga gtaataccca gcagcggggg gatttcgctt ttgttcgtgt gacctgggat 360
88 aaagcattaa agctgtttta acactcactc gatgaggtcc aaaccaagta cgggccatcg 420
89 ggcttacacg caggacaaac tggttggcgc gccacggggc aactgcattc cagcaccagc 480
90 catatgcagc gcgcggtggg gatgcacggt aattttgtga aaaaaatcgg cgactactcc 540
91 accggtgcag gccaaaccatt ctgccctatg tattaggtct aaccgaagta tatgcccaag 600
92 gcacctcttg gccactgatc ttagaaaaa gcaacacgat tgtgctgtgg tcaaacaatc 660
93 cttacaaaaa cctgcaagtg ggctggaacg ctgaaaccca tgaggccttt gcgtacctcg 720
94 cgcaattaaa agagaaggtc aaacagggta aaatccgctg gatcagtatc gacctgtgtg 780
95 tgactaaaac ccaggcttac cttggctgtg agcagttgta tgtgaacca cagactgacg 840
96 tgacgtgat gctggccatc gcccatgaga tgatcaccca aaagctacac gatgagaaat 900
97 tcatccaagg ttacagctta ggctttgaag agtttgtgcc ttacgtgatg ggcactaaag 960
98 atgggtatcg caaaaccctt gattgggcgg cgcctatctg cgggtgtgaa ccacacatta 1020
99 tccgcgatct ggccaaaacc ttagttaagg gccgtaccca aataatgatg ggttggtgta 1080
100 ttcagcgcca acaacacggt gagcaacctt actggatggc cgcagtaactt gcgaccatga 1140
101 taggccaaat cggcttacct ggcggtggta ttagctatgg tcaccactac tccagtattg 1200
102 gtgtgccggc gaccacagct gcagctccgg gcgctttccc acgtaactta gacgagaatc 1260
103 aaaaaccgct gtttgacagc acagacttta aaggcgcaag cagcacgatt cccgttgccc 1320
104 gctggattga tgcgattctc gaaccgggca aaaccattga tgccaacggt tctaaagtgg 1380
105 tatatcccga tattaagatg atgattttct cgggtaataa cccatggaac caccatcaag 1440
106 accgtaaccg catgaagcaa gccttccaaa agcttgaatg cgtgggtctct attgatgtga 1500

```

## RAW SEQUENCE LISTING

DATE: 10/08/2003

PATENT APPLICATION: US/10/088,117A

TIME: 14:51:42

Input Set : A:\05081004.app

Output Set: N:\CRF4\10082003\J088117A.raw

```

107 actggacagc cacttgctgt ttctccgaca tcgtgttgcc agcttgccacc acctatgagc 1560
108 gtaacgatat cgacgtgtac ggtgcctatg ccaaccgcg tattctggcg atgcagaaaa 1620
109 tggctcgagcc tttgtttgaa agcttgctctg actttgagat ctttactcgt tttgccgcgc 1680
110 tgctgggtaa agagaaagaa tacaccgcga atatgagcga aatggagtg atagaaaccc 1740
111 tctataacga gtgtaaagcc gctaaccgcg gtaagtatga gatgcctgac tttgccacct 1800
112 tctggaagca aggttatgta cactttggtg aaggcgaatt atggacgcgc cacgctgact 1860
113 ttagaaacga tcctgaaatc aaccattag gcacgccttc cggtttgatt gaaatcttca 1920
114 gtcgtaagat tgaacaattt ggctatgatg actgccaagg ccatcctatg tggatggaaa 1980
115 aagctgaacg cagccatggt ggtccaggtt cgaataaata tcctatgtgg ttgcaatctt 2040
116 gccatccgga tcaccgctta cactcacaaa tgtgtgagtc aaaggaatac cgcgaaacct 2100
117 acacagttaa tggctcgcaa cctgtgtata ttagccctga agatgctaaa acccggtgca 2160
118 ttaaagatgg cgatatcgtg cgggtcttta acgaccgagg tcaactgtta gctggcgag 2220
119 ttggtatcgga tcgtttccct aaagggtgtag tgcgaattca tgaaggtgca tggtagggcc 2280
120 cagtgggtaa agatggcagc gttgaaggcg gagccgaaat cgggtgccctg tgcagctatg 2340
121 gtgaccctaa taccctaacc ttagacattg gtacctctaa gttggctcaa gcttgctcag 2400
122 cctatacatg tctggttgag tttgaaaaat accaaggtaa agcacctaag gttagctcct 2460
123 tcgatggtcc tatcgaagtt gaaatc 2486
126 <210> SEQ ID NO: 3
127 <211> LENGTH: 2487
128 <212> TYPE: DNA
129 <213> ORGANISM: Shewanella c
131 <400> SEQUENCE: 3
132 atgaacagaa gagacttttt aaagggtatc gcctcatcct ctttcgttgt cttaggtggc 60
133 agctcagtg tagcgccctt aaatgcctta gccaaaacgg gcatcaatga agacgaatgg 120
134 ctaaccacag gttcacactt cggcgccctt aaaatgaagc gcaaaaacgg cgtcattgcc 180
135 gaagtgaaac ccttcgactt agataagtat ccaacggata tgattaacgg catccgcgac 240
136 atggtctaca atccatcgcg tgtacgttac cctatggtgc gcttagattt tttactcaaa 300
137 ggtcataaga gtaataccca tcaacggggt gatttccgct ttgttcgtgt aacatgggac 360
138 aaggcattaa cactgtttta gcattcatta gatgaagtcc aaaccaata cgtccatca 420
139 ggtctgcatg cgggtcaaac tgggtggcgc gccacgggtc aactgcattc cagcacgagt 480
140 catatgcaac gtgcggtggg gatgcacggc aactatgtga agaaaatcgg cgactactcc 540
141 acaggtgcag gccaaacaat tctgccctac gtggttaggt caaccgaagt gtagcccaa 600
142 ggcacttcat ggccgctgat cttagaacac agcgacacta tcgtgctctg gtcgaacgat 660
143 ccgtacaaga acctgcaagt gggttggaat gcggaiaacc atgaatcttt tgcttatctt 720
144 ggcaggttaa aagagaaagt gaagcaaggc aagatccgtg ttatcagtat cgaccctgtg 780
145 gtgactaaga cccaagccta tttgggctgt gagcaactct acgttaaccc acagacagac 840
146 gtgactttaa tgctggccat cgcccacgag atgatcagca aaaagctcta cgacgataaa 900
147 tttatccaag gctacagctt aggttttgaa gagtttgtgc cctatgtgat gggactaaa 960
148 gatggcgtag ccaaaacccc agaatgggcc gcgcctatct gtggtgttga agcccatgtt 1020
149 atccgcgact tggctaaaac cttagtcaag ggccgcactc agttcatgat gggctggtgt 1080
150 atccagcgcc agcaacacgg cgaacaaccc tattggatgg cggcggtact ggcgaccatg 1140
151 atcggccaaa tcggtctacc cgggtggtggc atcagttatg gtcaccaata ctcgagtatc 1200
152 ggcgtgcctt catcggtgac cgcggcgcca ggtgctttcc cccgtaactt ggacgaaaat 1260
153 caaaagccac tatttgatag ctcagacttc aaggcgcgca gcagcacaat tccggttgcc 1320
154 cgctggattg atcgattct cgaacctggg aaaaccattg atgctaacgg ctcgaaagt 1380
155 gtttatccc atatacaagat gatgattttc tcgggtaata atccttgga ccatcaccaa 1440
156 gacagaaacc gtatgaagca agccttccat aagcttgagt gtgtggtcac tgttgatgtg 1500
157 aactggacgg caacttgccg cttctcggt atcgctactac ccgcttgtag tacctatgag 1560
158 cgcaacgata tcgacgttta cggcgccctat gctaaccgcg gtatttttag catgcagaaa 1620

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/088,117A

DATE: 10/08/2003

TIME: 14:51:42

Input Set : A:\05081004.app

Output Set: N:\CRF4\10082003\J088117A.raw

```

159 atggttgagc cactgtttga tagcttgtcg gattttgaaa ttttcactcg ctttgccgcc 1680
160 gtacttggtg aagagaaaga atacaccgct aacatgggcg aaatggagtg gctagaaacc 1740
161 ctctataacg aatgtaaagc cgccaacgcg ggcaagtttg agatgcctga ctttgcgact 1800
162 ttctggaaac aaggttatgt gcatttttgt gacggtgaac tctggacgcg ccatgcagac 1860
163 tttagaaacg atcctgaaat caatccacta ggacgcctt caggtttgat tgaaatcttt 1920
164 agccgtaaga ttgatcaatt cggttacgat gactgtaaag gtcacccaac ttggatggag 1980
165 aaaaccgagc gtagccatgg cggccctggg tctgacaagc atccgatttg gttgcagtca 2040
166 tgccaccagc acaaacgctt acactcgcaa atgtgtgagt cgcgagaata ccgcgagacc 2100
167 tacgcagtca atggccgtga gcctgtgtat atcagccctg tcgacgcaaa agcgcgcggc 2160
168 ataaaagatg gcgatatagt gcgagctttt aacgaccgtg gccactgtt ggcggttgca 2220
169 gtggtatcgg acaacttccc tactgggtat gtgcggattc acgaaggcgc atggtatggg 2280
170 ccagtaggta aagatggtag cactgaaggt ggggctgaag tcggcgccct gtgcagttat 2340
171 ggcgatccta acacctcac tttagacata ggcacatcta aacttgccca agcttgctca 2400
172 gcctatactt gcttagtcga gtttgagaaa taccaaggca aagtgcctaa ggtcagctcc 2460
173 ttcgatggcc ctatcgaagt cgaaatc 2487
176 <210> SEQ ID NO: 4
177 <211> LENGTH: 1080
178 <212> TYPE: DNA
179 <213> ORGANISM: Photobacterium phosphoreum
181 <400> SEQUENCE: 4
182 acaatactga aagattgtaa gacattgata tgggtgtcaa atgatccgat taaaaacagt 60
183 caggttggct ggcagtgtga gactcatggt tcttatgagt attatgcgca attaaagcag 120
184 aaggctgcag atggtgggat ccgtatgata tcggtcgatc ctgtagtgtc gaaatcgcaa 180
185 aaatatttta actgtgagca ccaatacgtc aatcctcaaa ctgacgttcc tttcatgctt 240
186 gctattgcgc atacattgta taaagaagat ctgtacgata acaatttct ggaaacttac 300
187 acttttaggt tcaatgaatt cttgccttac ttattgggta caggcaaaga taaaatagcc 360
188 aaaacgccag aatgggcaga gccaatgtgt ggcgttaaag cagaggctat tcgagaattt 420
189 gctcgcggat tagttaaaaa ccgtacgatg ataatgtttg gttgggctgt acagcgtcaa 480
190 caacacggtg agcagcctta ttggatggga gcagtgtctg cttcgatgtt aggccaaata 540
191 ggcttacctg gtggagggat ttcctattct cacttttaca gtggcgttgg gttacctttc 600
192 agtactgcag ctgggcccgg gggatttccg cgtaatgttg atgaaggcca acagccgatt 660
193 tggaataata acgattttta aggctacagt tcgacaattc cggtcgcaag atggattgat 720
194 gcgatcatgg aaccaggtaa aaaaattcaa tataacggcg ctaatgtggt gttgcctgat 780
195 attaatgatg tgggtctttg tggttgaat ccgtggaatc atcatcaaca acgtaatcgt 840
196 atgaaacaag catttagaaa gctgcaaacc gtgggttaata ttgattatac atggacacca 900
197 acctgtcgtt tttccgatat tgtattacct gcttgtaacc aatttgagcg tagtgattta 960
198 gatcaatatg gtacttattc aactagcggg attttagcga tgcataagct aattgatccg 1020
199 ctttatcaat caaaaacaga ctttcagata tttactgaat taaccgaacg ctttgggaaa 1080
202 <210> SEQ ID NO: 5
203 <211> LENGTH: 392
204 <212> TYPE: PRT
205 <213> ORGANISM: Shewanella massilia
207 <400> SEQUENCE: 5
208 Met Lys Trp Leu Thr Asn Leu Trp Arg Thr Leu Asn Lys Pro Thr Lys
209 1 5 10 15
211 Ala Leu Thr Leu Gly Ala Val Ser Ile Ser Ala Phe Ile Met Gly Ile
212 20 25 30
214 Ile Phe Trp Gly Gly Phe Asn Thr Ala Leu Glu Ala Thr Asn Thr Glu
215 35 40 45

```

## RAW SEQUENCE LISTING

DATE: 10/08/2003

PATENT APPLICATION: US/10/088,117A

TIME: 14:51:42

Input Set : A:\05081004.app

Output Set: N:\CRF4\10082003\J088117A.raw

```

217 Ala Phe Cys Ile Ser Cys His Ser Met Glu Ser Lys Pro Tyr Gln Glu
218      50      55      60
220 Leu Gln Glu Thr Val His Trp Ser Asn His Phe Gly Val Arg Ala Thr
221 65      70      75      80
223 Cys Pro Asp Cys His Val Pro His Asn Trp Ser Arg Lys Ile Ala Arg
224      85      90      95
226 Lys Met Glu Ala Ser His Asp Val Trp Gly Trp Leu Phe Asn Thr Val
227      100      105      110
229 Asn Thr Pro Glu Lys Phe Glu Ala Lys Arg Leu Glu Met Ala Ser Arg
230      115      120      125
232 Glu Trp Lys Arg Phe Asp Arg Asp Asn Ser Leu Ala Cys Lys Asn Cys
233      130      135      140
235 His Asn Tyr Asn Ser Met Lys Trp Glu Ala Met Ser Pro Leu Ala Gln
236 145      150      155      160
238 Lys Gln Met Lys Arg Ala Ala Glu Ile Asp Gln Ser Cys Ile Asp Cys
239      165      170      175
241 His Lys Gly Ile Ala His His Leu Pro Glu Met Gly Thr Ala Arg Ala
242      180      185      190
244 Pro Glu Leu Ile Ala Glu Val Gly Ala Gly Val Ser Ser Val Glu Thr
245      195      200      205
247 Asn Gln Thr Tyr Tyr Ser Ala Leu Thr Lys Pro Leu Phe Phe Thr Asp
248      210      215      220
250 Lys Gly Asp Val Glu Ala Gly Thr Leu Asn Val Ala Thr Lys Val Lys
251 225      230      235      240
253 Val Leu Glu Thr Gln Gly Lys Arg Ile Lys Ile Gly Ile Asp Gly Trp
254      245      250      255
256 Arg Lys Lys Ile Gly Ala Gly Arg Val Ile Tyr Met Asp Phe Gly Val
257      260      265      270
259 Asn Ile Leu Ser Ala Gln Leu Thr Lys Asp Ala Ala Glu Thr Gly Gly
260      275      280      285
262 Val Ile Gln Thr Phe Glu Glu Lys Glu Asp Pro Met Thr Gly Leu Lys
263      290      295      300
265 Trp Gln Arg Ile Glu Ala Gln Ile Trp Thr Asp Lys Asp Tyr Leu Leu
266 305      310      315      320
268 Thr Glu Leu Gln Pro Leu Trp Gly Tyr Ala Arg Asp Thr Phe Arg Ser
269      325      330      335
271 Ser Cys Ser Val Cys His Thr Gln Pro Asp Glu Ala His Phe Asp Ala
272      340      345      350
274 Asn Thr Trp Pro Gly Met Phe Gln Gly Met Leu Ala Phe Val Asn Met
275      355      360      365
277 Asp Gln Asp Thr Gln Ala Leu Val Gln Lys Tyr Leu Gln Glu His Ser
278      370      375      380
280 Ser Thr Phe Val Lys Lys Glu His
281 385      390
284 <210> SEQ ID NO: 6
285 <211> LENGTH: 2523
286 <212> TYPE: DNA
287 <213> ORGANISM: Rhodobacter sphaeroides
289 <400> SEQUENCE: 6

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/088,117A

DATE: 10/08/2003  
TIME: 14:51:43

Input Set : A:\05081004.app  
Output Set: N:\CRF4\10082003\J088117A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:22; N Pos. 9  
Seq#:24; N Pos. 3,6  
Seq#:27; N Pos. 3,15

. . . .  
  
**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/088,117A

DATE: 10/08/2003

TIME: 14:51:43

Input Set : A:\05081004.app

Output Set: N:\CRF4\10082003\J088117A.raw

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:752 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22 after pos.:0  
L:786 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24 after pos.:0  
L:832 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27 after pos.:0